BULL. AUSTRAL. MATH. SOC. 05B40
VOL. 53 (1995) [267–269]

SET COVERING NUMBER FOR A FINITE SET

H.-C. CHANG AND N. PRABHU

Given a finite set S of cardinality N, the minimum number of j-subsets of S needed to cover all the r-subsets of S is called the covering number $C(N, j, r)$. While Erdös and Hanani's conjecture that $\lim_{N\to\infty} (C(N, j, r))/\left(\binom{N}{r}/\binom{j}{r}\right) = 1$ was proved by Rödl, no nontrivial upper bound for $C(N, j, r)$ was known for finite N. In this note we obtain a nontrivial upper bound by showing that for finite N, $C(N, j, r) \leqslant \left(\binom{N}{r}/\binom{j}{r}\right) \ln \binom{N}{r}$.

Let S be a set with N elements. If $J_1, \ldots, J_k$ are j-subsets of S, (that is, subsets of S of cardinality j) then $\mathcal{J} = \{J_1, \ldots, J_k\}$ is called a k-collection of j-subset of S. Further, the k-collection $\mathcal{J}$ is said to cover all the r-subsets of S, if for every r-subset R of S, there is some $J_i \in \mathcal{J}$, $1 \leqslant i \leqslant k$, such that $R \subset J_i$. The set covering number $C(N, j, r)$, $j > r$, is the smallest integer k such that some k-collection of j-subsets of S covers all the r-subsets of S. Clearly

$$C(N, j, r) \geqslant \frac{\binom{N}{r}}{\binom{j}{r}},$$

and in 1963, Erdös and Hanani [1] conjectured that

$$\lim_{N\to\infty} \frac{C(N, j, r)}{\binom{N}{r}/\binom{j}{r}} = 1.$$

Erdös and Hanani's conjecture was proved by Rödl [2] in 1985. However, for finite N a nontrivial upper bound for $C(N, j, r)$ was not known. In this note, using probabilistic arguments we obtain a nontrivial upper bound for $C(N, j, r)$ by proving:

Received 1 June 1995
Supported in part by NSF Research Initiation Award

Theorem 1. *For $N \geqslant j > r$,*

$$C(N, j, r) \leqslant \frac{\binom{N}{r}}{\binom{j}{r}} \ln \binom{N}{r}.$$

Proof: Let S be a finite set of cardinality N. Let $\mathcal{J} = \{J_1, \dots, J_k\}$ be a k-collection of j-subsets of S, where J_i, $1 \leqslant i \leqslant k$ are chosen randomly and independently from among the $\binom{N}{j}$ j-subsets of S; $J_1, \dots, J_k$ need not all be distinct.

Consider an r-subset R of S. Then for any j-subset $\widetilde{J}$ of S, the probability that $R \subset \widetilde{J}$ is

$$P[R \subset \widetilde{J}] = \frac{\binom{N-r}{j-r}}{\binom{N}{j}} = \frac{(N-r)!\,(N-j)!\,j!}{N!\,(N-j)!\,(j-r)!} = \frac{(N-r)!\,j!\,r!}{N!(j-r)!\,r!}$$
$$= \frac{\binom{j}{r}}{\binom{N}{r}}.$$

Therefore

$$P[R \not\subset \widetilde{J}] = 1 - \frac{\binom{j}{r}}{\binom{N}{r}}.$$

The probability that none of the randomly chosen j-subsets of $\mathcal{J}$, namely $J_1, \dots, J_k$, contains R is hence

$$R[(R \not\subset J_1) \wedge \dots \wedge (R \not\subset J_k)] = \left(1 - \frac{\binom{j}{r}}{\binom{N}{r}}\right)^k.$$

Label the r-subsets of S, $R_1, \dots, R_{\binom{N}{r}}$. Let A_i be the event that R_i does not belong to any of the j-subsets of $\mathcal{J}$, $1 \leqslant i \leqslant \binom{N}{r}$. Then the probability that a randomly chosen k-collection of j-subsets $\mathcal{J}$ does not contain at least one r-subset is

$$R[A_1 \vee \dots \vee A_{\binom{N}{r}}] \leqslant \binom{N}{r} \left(1 - \frac{\binom{j}{r}}{\binom{N}{r}}\right)^k.$$

Therefore if $P[A_1 \vee \ldots \vee A_{\binom{N}{r}}] < 1$, then $P[\overline{A}_1 \wedge \ldots \wedge \overline{A}_{\binom{N}{r}}] > 0$ and hence some k-collection of j-subsets must cover all the r-subsets of S. $\overline{A}_i$ denotes the complement of the event A_i, $1 \leqslant i \leqslant \binom{N}{r}$. Thus we want

$$\binom{N}{r}\left(1-\frac{\binom{j}{r}}{\binom{N}{r}}\right)^k < 1,$$

or

$$k\ln\left(1-\frac{\binom{j}{r}}{\binom{N}{r}}\right)+\ln\binom{N}{r} < 0,$$

or

$$k > \frac{-\ln\binom{N}{r}}{\ln\left(1-\binom{j}{r}/\binom{N}{r}\right)}. \tag{1}$$

The direction of the inequality is changed in equation (1) since, for $j < N$, $\binom{j}{r} < \binom{N}{r}$ and hence $\ln\left(1-\frac{\binom{j}{r}}{\binom{N}{r}}\right) < 0$.

For $0 < x < 1$, $-\ln(1-x) \geqslant x$, since $f(x) = e^{-x} - 1 + x \geqslant 0$ for $0 < x < 1$. ($f(0) = 0$ and $f^1(x) > 0$ for $0 < x < 1$.) Hence

$$k > \frac{\ln\binom{N}{r}}{\binom{j}{r}/\binom{N}{r}} \Rightarrow k > \frac{-\ln\binom{N}{r}}{\ln\left(1-\binom{j}{r}/\binom{N}{r}\right)}.$$

Therefore, for any $k > \left(\binom{N}{r}/\binom{j}{r}\ln\binom{N}{r}i\right)$, $P[\overline{A}_1 \wedge \ldots \wedge \overline{A}_{\binom{N}{r}}] > 0$ and hence there exists a k-collection of j-subsets of S which covers all the r-substs of S. $\square$

References

[1] P. Erdös and H. Hanani, 'On a limit theorem in combinatorial analysis', *Publ. Math. Debrecen* **10** (1963), 10–13.

[2] V. Rödl, 'On a packing and covering problem', *European J. Combin.* **5** (1985), 69–78.

Department of Mathematics
Purdue University MGL 1303
West Lafayette IN 47907
United States of America